

SUMMARY OF THE FIGURES

[0046] For better understanding of the present application, reference will now be made by way of example to the accompanying drawings in which:

[0047] FIG. 1 shows schematically a multi-user free-view-point service sharing system which may encompass embodiments of the application;

[0048] FIG. 2 shows schematically an apparatus suitable for being employed in embodiments of the application;

[0049] FIG. 3 shows schematically an example content co-ordinating apparatus according to some embodiments;

[0050] FIG. 4 shows a flow diagram of the operation of the example content co-ordinating apparatus shown in FIG. 3 according to some embodiments;

[0051] FIG. 5 shows an example audio segment; and

[0052] FIGS. 6 to 9 show audio alignment examples according to some embodiments.

EMBODIMENTS OF THE APPLICATION

[0053] The following describes in further detail suitable apparatus and possible mechanism for the provision of effective audio signal capture sharing. In the following examples, audio signals and audio capture signals are described. However it would be appreciated that in some embodiments the audio signal/audio capture is a part of an audio-video system.

[0054] The concept of this application is related to assisting in the production of immersive person-to-person communication and can include video. It would be understood that the space within which the devices record the audio signal can be arbitrarily positioned within an event space. The captured signals as described herein are transmitted or alternatively stored for later consumption where the end user can select the listening point based on their preference from the reconstructed audio space. The rendering part then can provide one or more down mixed signals from which the multiple recordings that correspond to the selective listening point. It would be understood that each recording device can record the event seen and upload or upstream the recorded content. The uploaded or upstream process can include implicitly positioning information about where the content is being recorded.

[0055] Furthermore an audio scene can be defined as a region or area within which a device or recording apparatus effectively captures the same audio signal. Recording apparatus operating within an audio scene and forwarding the captured or recorded audio signals or content to a co-ordinating or management apparatus effectively transmit many copies of the same or very similar audio signal. The redundancy of many devices capturing the same audio signal permits the effective sharing of the audio recording or capture operation.

[0056] Content or audio signal discontinuities can occur, especially when the recorded content is uploaded to the content server after some time the recording has taken place that the uploaded content represents an edited version rather than the actual recorded content. For example the user can edit any recorded content before uploading the content to the content server. The editing can for example involve removing unwanted segments from the original recording. The signal discontinuity can create significant challenges to the content server as typically an implicit assumption is made that the uploaded content represents the audio signal or clip from a continuous timeline. Where segments are removed (or added) after recording has ended then the continuity assumption or condition no longer holds for the particular content.

[0057] Furthermore to be able to jointly utilize the multi-user recorded content for various media rendering methods, such as audio mixing from multiple users and video view switching from one user to the other, the content between different users must employ a 'common' time or timeline. Furthermore, the common timeline should be constructed such that the content from different devices or apparatus shares the same event space. For example users and their apparatus or devices may move in and out of a defined audio event space during recording or capturing resulting in a situation where there may be time periods when some apparatus do not share the same event space even though they share the same timeline. Furthermore depending on the event venue there may exist multiple event spaces that are not correlated. For example an event venue with different rooms and/or floors can result in multiple event spaces from the content capturing and rendering point of view.

[0058] The concept as described herein in embodiments is to analyse and segment the recorded or captured content from an event venue into different event spaces. This invention outlines method for creating event spaces from multi-user captured content. The concept can further be summarized according to the following steps:

[0059] Classifying recorded or captured media content to generate media segments associated with a defined class

[0060] Applying analysis to media segments based the associated class

[0061] Determining similarities between segments of different user/apparatus media

[0062] Creating event spaces based on similarity status for different user/apparatus media

[0063] In some embodiments the classification comprises at least 2 classes, for example a music class and a non-music class. Furthermore in some embodiments the classes can furthermore be sub-divided into subclasses of which the subclasses are grouped, for example the music class can be divided into a music-classical, and music-rock sub-classes.

[0064] In some embodiments the media analysis is applied to each class present in the segment from different user media.

[0065] In some embodiments the audio domain properties are used to provide event space separation resulting fast and computationally efficient operation.

[0066] With respect to FIG. 1 an overview of a suitable system within which embodiments of the application can be located is shown. The audio space 1 can have located within it at least one recording or capturing device or apparatus 19 which are arbitrarily positioned within the audio space to record suitable audio scenes. The apparatus 19 shown in FIG. 1 are represented as microphones with a polar gain pattern 101 showing the directional audio capture gain associated with each apparatus. The apparatus 19 in FIG. 1 are shown such that some of the apparatus are capable of attempting to capture the audio scene or activity 103 within the audio space. The activity 103 can be any event the user of the apparatus wishes to capture. For example the event could be a music event or audio of a "news worthy" event. The apparatus 19 although being shown having a directional microphone gain pattern 101 would be appreciated that in some embodiments the microphone or microphone array of the recording apparatus 19 has a omnidirectional gain or different gain profile to that shown in FIG. 1.

[0067] Each recording apparatus 19 can in some embodiments transmit or alternatively store for later consumption the